In Vitro Selection of a DNA-Templated Small-Molecule Library Reveals a Class of Macrocyclic Kinase Inhibitors

J|A|C|S ARTICLES

Published on Web 08/03/2010

In Vitro Selection of a DNA-Templated Small-Molecule Library Reveals a Class of Macrocyclic Kinase Inhibitors

Ralph E. Kleiner, Christoph E. Dumelin, Gerald C. Tiu, Kaori Sakurai, and David R. Liu*

Department of Chemistry and Chemical Biology and the Howard Hughes Medical Institute, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

Received June 4, 2010; E-mail: drliu@fas.harvard.edu

Abstract: DNA-templated organic synthesis enables the translation of DNA sequences into synthetic small-molecule libraries suitable for *in vitro* selection. Previously, we described the DNA-templated multistep synthesis of a 13 824-membered small-molecule macrocycle library. Here, we report the discovery of small molecules that modulate the activity of kinase enzymes through the *in vitro* selection of this DNA-templated small-molecule macrocycle library against 36 biomedically relevant protein targets. DNA encoding selection survivors was amplified by PCR and identified by ultra-high-throughput DNA sequencing. Macrocycles corresponding to DNA sequences enriched upon selection against several protein kinases were synthesized on a multimilligram scale. *In vitro* assays revealed that these macrocycles inhibit (or activate) the kinases against which they were selected with IC_{50} values as low as 680 nM. We characterized in depth a family of macrocycles enriched upon selection against Src kinase, and showed that inhibition was highly dependent on the identity of macrocycle building blocks as well as on backbone conformation. Two macrocycles in this family exhibited unusually strong Src inhibition selectivity even among kinases closely related to Src. One macrocycle was found to activate, rather than inhibit, its target kinase, VEGFR2. Taken together, these results establish the use of DNA-templated synthesis and *in vitro* selection to discover small molecules that modulate enzyme activities, and also reveal a new scaffold for selective ATP-competitive kinase inhibition.

Introduction

The discovery of small molecules capable of selectively modulating the activity of biological targets remains a central challenge of chemistry and chemical biology. Such small molecules are commonly discovered through combinatorial[1,2] or diversity-oriented[3] synthesis and high-throughput screening[4] (HTS). In contrast, functional molecules emerge in nature through iterated cycles of translation, selection, and amplification with mutation.[5−8] While scientists have applied components of biological evolution to generate DNA, RNA, and protein molecules with tailor-made catalytic or binding properties, this approach has traditionally been restricted to molecules whose structures are compatible with biosynthetic machinery.[9−16] Our group has developed DNA-templated organic synthesis as a method for translating DNA sequences into synthetic small molecules[17−25] and synthetic polymers[26−28] that can be sub-

(1) Lam, K. S.; Lebl, M.; Krchnak, V. *Chem. Rev.* **1997**, *97*, 411–448.
(2) Pirrung, M. C. *Chem. Rev.* **1997**, *97*, 473–488.
(3) Tan, D. S. *Nat. Chem. Biol.* **2005**, *1*, 74–84.
(4) Walters, W. P.; Namchuk, M. *Nat. Rev. Drug Discovery* **2003**, *2*, 259–266.
(5) Joyce, G. F. *Annu. Rev. Biochem.* **2004**, *73*, 791–836.
(6) Lin, H.; Cornish, V. W. *Angew. Chem., Int. Ed.* **2002**, *41*, 4402–4425.
(7) Taylor, S. V.; Kast, P.; Hilvert, D. *Angew. Chem., Int. Ed.* **2001**, *40*, 3310–3335.
(8) Wilson, D. S.; Szostak, J. W. *Annu. Rev. Biochem.* **1999**, *68*, 611–648.
(9) Dewey, T. M.; Mundt, A. A.; Crouch, G. J.; Zyzniewski, M. C.; Eaton, B. E. *J. Am. Chem. Soc.* **1995**, *117*, 8474–8475.
(10) Forster, A. C.; Tan, Z.; Tan, M. N.; Nalam, H.; Lin, H.; Qu, H.; Cornish, V. W.; Blacklow, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6353–6357.
(11) Horhota, A.; Zou, K.; Ichida, J. K.; Yu, B.; McLaughlin, L. W.; Szostak, J. W.; Chaput, J. C. *J. Am. Chem. Soc.* **2005**, *127*, 7427–7434.
(12) Josephson, K.; Hartman, M. C. T.; Szostak, J. W. *J. Am. Chem. Soc.* **2005**, *127*, 11727–11735.
(13) Latham, J. A.; Johnson, R.; Toole, J. J. *Nucleic Acids Res.* **1994**, *22*, 2817–2822.
(14) Lee, S. E.; Sidorov, A.; Gourlain, T.; Thorpe, S. J.; Brazier, J. A.; Dickman, M. J.; Hornby, D. P.; Grasby, J. A.; Williams, D. M. *Nucleic Acids Res.* **2001**, *29*, 1565–1573.
(15) Matulic-Adamic, J.; Daniher, A. T.; Karpeisky, A.; Haeberli, P.; Sweedler, D.; Beigelman, L. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1299–1302.
(16) Perrin, D. M.; Garestier, T.; Helène, C. *J. Am. Chem. Soc.* **2001**, *123*, 1556–1663.
(17) Gartner, Z. J.; Liu, D. R. *J. Am. Chem. Soc.* **2001**, *123*, 6961–6963.
(18) Gartner, Z. J.; Kanan, M. W.; Liu, D. R. *Angew. Chem., Int. Ed.* **2002**, *41*, 1796–1800.
(19) Gartner, Z. J.; Kanan, M. W.; Liu, D. R. *J. Am. Chem. Soc.* **2002**, *124*, 10304–10306.
(20) Gartner, Z. J.; Tse, B. N.; Grubina, R.; Doyon, J. B.; Snyder, T. M.; Liu, D. R. *Science* **2004**, *305*, 1601–1605.
(21) Calderone, C. T.; Liu, D. R. *Angew. Chem., Int. Ed.* **2005**, *44*, 7383–7386.
(22) Snyder, T. M.; Liu, D. R. *Angew. Chem., Int. Ed.* **2005**, *44*, 7379–7382.
(23) Kanan, M. W.; Rozenman, M. M.; Sakurai, K.; Snyder, T. M.; Liu, D. R. *Nature* **2004**, *431*, 545–549.
(24) Tse, B. N.; Snyder, T. M.; Shen, Y. S.; Liu, D. R. *J. Am. Chem. Soc.* **2008**, *130*, 15611–15626.
(25) Li, X.; Liu, D. R. *Angew. Chem., Int. Ed.* **2004**, *43*, 4848–4870.

jected to *in vitro* selection for desired properties.[17,20,23,28,29] Several related approaches to generate and evaluate DNA-encoded small-molecule libraries have also been used successfully in academic[30−38] and industrial settings.[39,40]

Macrocycles are particularly attractive candidates for the discovery of biologically active small molecules because their rigid scaffolds can decrease the entropic cost of target binding and limit access to nonbinding conformations, resulting in higher affinity and greater binding specificity than their corresponding linear counterparts.[41] In addition, macrocyclic peptide-like structures can offer advantages for applications in cell culture and *in vivo* over their linear analogues, since they can possess higher bioavailability, membrane permeability, and resistance to *in vivo* degradation.[41] While synthesizing macrocyclic structures especially in a library format can be challenging,[42,43] we speculated that features of DNA-templated synthesis including compatibility with aqueous solvents, extremely low (nM) reactant concentrations, and the ability of base pairing to hold together relevant reactants at high effective molarities would promote efficient macrocylization. Indeed, these features enabled the DNA-templated synthesis and model *in vitro* selection of a pilot library of 65 macrocycles.[20] Subsequent advances in DNA template design and DNA-templated synthesis methods enabled the preparation and characterization of a larger 13 824-membered DNA-templated macrocycle library.[24]

Here, we report the discovery and characterization of selective inhibitors of protein kinases from the *in vitro* selection of the 13 824-membered DNA-templated macrocycle library against a panel of therapeutically relevant protein targets. In contrast to typical HTS technologies, selections enable the simultaneous evaluation of small-molecule libraries in one pot regardless of library size, obviating the significant time and infrastructure demands of screening. Moreover, the simplicity of these selections enables large numbers of them to be executed in parallel by a single researcher. The *in vitro* selection effort summarized in this work represents the evaluation of ~497 000 potential protein-small molecule interactions by a single researcher yet only required a modest time investment and simple equipment. To facilitate the analysis of such a large number of selection outcomes in a cost-effective and efficient manner, we used PCR-installed DNA barcodes in conjunction with ultra-high-throughput ("deep") DNA sequencing.

The Src kinase inhibitors discovered through this approach represent, to our knowledge, the first examples of synthetic peptidic macrocycles that inhibit protein kinase activity in an ATP-competitive manner. Some of the Src-inhibiting macrocycles exhibited unusual selectivity for Src when screened against a representative panel of human protein kinases. We also discovered macrocycles that activate VEGFR2 kinase and that inhibit Akt3, MAPKAPK2, p38α, and Pim1 kinases. Together, these results demonstrate that DNA-templated library synthesis coupled with *in vitro* selection can lead to the discovery of protein-binding synthetic small molecules that include enzyme inhibitors and activators. These results also reveal two novel and synthetically versatile scaffolds for the selective inhibition of Src-family protein kinases.

Results

Previous work in our group has established that DNA-linked small molecules with protein binding affinity can be enriched from complex mixtures containing predominantly nonbinding DNA-linked small molecules,[20,29] and that a library of 13 824 DNA templates can be translated into a corresponding library of synthetic macrocycles through three DNA-templated synthesis reactions.[24] In this work, we integrate these capabilities into a broad effort to discover and characterize synthetic small-molecule macrocycles that inhibit or activate protein targets of interest.

Validating the *in Vitro* Selection of DNA-Linked Small Molecules That Bind Proteins Fused to GST. Prior to undertaking library selections, we first validated a general and efficient *in vitro* selection strategy that would support selections against many different protein targets. We sought to improve our earlier *in vitro* selection efforts that used proteins covalently bound to beads[29] by implementing a strategy that would not perturb the covalent structure of the target protein upon immobilization. Glutathione-*S*-transferase (GST) fusions of many biomedically relevant proteins are readily available, typically retain native function, and can be immobilized in a defined way using glutathione-linked resin. Furthermore, library members that bind targets presented in this manner can be specifically eluted under mild conditions with free glutathione. This elution strategy should minimize the enrichment of library members that bind molecules other than the target of interest, such as resins or resin-bound linkers.

We tested the ability of immobilized GST-fusion proteins to enrich known small-molecule ligands covalently linked to DNA oligonucleotides in several mock selections using GST-tagged MDM2, FKBP, Bcl-xL, cRaf-1, and HSP90 as protein targets. In brief, a known small-molecule ligand linked to DNA (the positive control) was combined with a 100- to 10 000-fold molar excess of a nonbinding DNA sequence (the negative control). The negative and positive control sequences were of identical length and shared common PCR primer-binding sequences. The mixture was incubated with the corresponding protein target immobilized on glutathione-linked beads. Following several

(26) Kleiner, R. E.; Brudno, Y.; Birnbau, M. E.; Liu, D. R. *J. Am. Chem. Soc.* **2008**, *130*, 4646–4659.
(27) Rosenbaum, D. M.; Liu, D. R. *J. Am. Chem. Soc.* **2003**, *125*, 13924–13925.
(28) Brudno, Y.; Birnbaum, M. E.; Kleiner, R. E.; Liu, D. R. *Nat. Chem. Biol.* **2010**, *6*, 148–155.
(29) Doyon, J. B.; Snyder, T. M.; Liu, D. R. *J. Am. Chem. Soc.* **2003**, *125*, 12372–12373.
(30) Dumelin, C. E.; Scheuermann, J.; Melkko, S.; Neri, D. *Bioconjugate Chem.* **2006**, *17*, 366–370.
(31) Halpin, D. R.; Harbury, P. B. *PLoS Biol.* **2004**, *2*, 1015–1021.
(32) Halpin, D. R.; Harbury, P. B. *PLoS Biol.* **2004**, *2*, 1022–1030.
(33) Halpin, D. R.; Lee, J. A.; Wrenn, S. J.; Harbury, P. B. *PLoS Biol.* **2004**, *2*, 1031–1038.
(34) Melkko, S.; Scheuermann, J.; Dumelin, C. E.; Neri, D. *Nat. Biotechnol.* **2004**, *22*, 568–574.
(35) Melkko, S.; Zhang, Y.; Dumelin, C. E.; Scheuermann, J.; Neri, D. *Angew. Chem., Int. Ed.* **2007**, *46*, 4671–4674.
(36) Scheuermann, J.; Dumelin, C. E.; Melkko, S.; Zhang, Y.; Mannocci, L.; Jaggi, M.; Sobek, J.; Neri, D. *Bioconjugate Chem.* **2008**, *19*, 778–785.
(37) Wrenn, S. J.; Weisinger, R. M.; Halpin, D. R.; Harbury, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 13137–13143.
(38) Buller, F.; Steiner, M.; Scheuermann, J.; Mannocci, L.; Nissen, I.; Kohler, M.; Beisel, C.; Neri, D. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4188–4192.
(39) Clark, M. A.; et al. *Nat. Chem. Biol.* **2009**, *5*, 647–654.
(40) Hansen, M. H.; Blakskjaer, P.; Petersen, L. K.; Hansen, T. H.; Hojfeldt, J. W.; Gothelf, K. V.; Hansen, N. J. V. *J. Am. Chem. Soc.* **2009**, *131*, 1322–1327.
(41) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. *Nat. Rev. Drug Discovery* **2008**, *7*, 608–624.
(42) Hruby, V.; Ahn, J. M.; Liao, S. *Curr. Opin. Chem. Biol.* **1997**, *1*, 114–119.
(43) Spatola, A. F.; Romanovskis, P. In *Combinatorial Peptide and Nonpeptide Libraries: A Handbook*; Jung, G., Ed.; VCH Publishers: New York, 1996; pp 327−347.

successive washes with buffer, molecules that remained bound to the resin were eluted with free glutathione and subjected to PCR amplification or carried through to another round of selection. The ratio of positive control DNA to nonbinding DNA was measured using a restriction endonuclease[29] or by qPCR.[44−46] After only one round of selection, we observed substantial (∼10- to 1000-fold) enrichment of the positive control-encoding DNA sequence against all five GST-linked protein targets (Figure S1 and Table S1). These results confirmed our ability to enrich DNA-linked small-molecule protein ligands using immobilized GST-tagged proteins.

***In Vitro* Selection of 13 824 DNA-Templated Small-Molecule Macrocycles.** As previously described, the DNA-templated small-molecule macrocycle library used in this *in vitro* selection effort was generated in three DNA-templated reactions followed by a nontemplated Wittig macrocyclization reaction to generate the final product (Figure 1a).[20,24] Each DNA-templated library synthesis step relied on a set of 12 DNA-linked building blocks comprising both natural and non-natural amino acids (Figure 1b).[24] These 36 building blocks together with eight variable diamino-acid scaffolds (Figure 1b) result in a theoretical diversity of 13 824 DNA-templated macrocycles. The fidelity of DNA-templated macrocycle library synthesis was extensively characterized by high-resolution LC-MS analysis of sublibraries and by gel electrophoresis and LC-MS analysis of reactions containing all possible building blocks and scaffolds as previously described.[24]

We chose a diverse set of biomedically relevant protein targets including protein kinases, protein phosphatases, small GTPases, PDZ domains, SH2 domains, nuclear receptors, and antiapoptotic proteins for *in vitro* selection (Table 1). Multiple representatives were chosen from each protein class so that we could compare selection results within protein families. In total, we performed a single round of *in vitro* selection against 36 protein targets. Additionally, we also performed a selection for binding to GST (with no fused target protein) immobilized on glutathione-linked resin. Although successive rounds of selection (with or without DNA amplification and translation between rounds) can multiply net enrichment factors,[29] we hypothesized that recent advances in ultra-high-throughput DNA sequencing would enable even modest enrichment factors to be robustly detected, and the findings in this work are the result of a single round of selection on each protein target.

Following *in vitro* selection, DNA templates linked to target-binding macrocycles were amplified in PCR reactions using primers containing "barcode" sequences specific to each selection (Figure 2). We also amplified DNA templates from eight aliquots of the unselected library to determine the initial abundance of each library member. A set of 12 five-base barcodes used in pairwise combinations provided sufficient encoding complexity for one round of selection against 36 protein targets, a control selection against immobilized GST, and eight preselection library samples (Table S2). These barcodes were designed to differ by at least three bases from one another so that common DNA sequencing errors would not preclude barcode assignment. After affinity selection and PCR amplification, barcoded samples were combined in equimolar amounts and submitted for ultra-high-throughput DNA sequencing[47,48] as a single sample.

Ultra-High-Throughput DNA Sequencing and Analysis of Selection Results. To rapidly identify the DNA-linked macrocycles that survived target affinity selection, we used Solexa (Illumina) DNA sequencing technology.[49] Because standard Solexa sequencing read lengths at the time of this experiment were shorter than the length of our barcoded library templates (84 base pairs), the paired-end sequencing method, which provides both forward and reverse sequencing reads for each template, was used. We created a PERL computer program to analyze the large amount of data emerging from high-throughput sequencing. First, forward and reverse sequence reads were merged to give complete template sequences. Each complete DNA sequence was converted into a combination of a barcode (selection identity) and three codons (building block identities). Template sequences were then binned by barcode to deconvolute the results of each selection. The abundance fraction of each library member was determined within each selection by counting the number of corresponding DNA sequence occurrences and dividing by the total number of interpretable sequence reads for that selection. Finally, enrichment factors for each library member were computed by dividing postselection abundance fraction by preselection abundance fraction. In total, we received 25.6 million paired forward and reverse sequence reads from two Solexa sequencing runs. Of these, all four codons (which require five out of six correct bases per codon for the building blocks, and three out of three correct bases per codon for the scaffold[24]) and the selection barcode (which requires four out of five correct bases) could be conclusively assigned for 12.4 million sequences. As a result, we obtained ∼200 000 sequence reads for each selection, and 1.7 million sequence reads of the preselection library.

Although 100% of the 13 824 possible library DNA sequences were observed at least once, due to the varying efficiencies of library member synthesis,[24] library members in the preselection pool were not expected to be present in equal abundance. Since underrepresented library members are more prone to inaccurate enrichment factors arising from limited sampling, we evaluated selection results using scatter plots of enrichment factor versus sequence counts in order to better identify unusually high enrichment factors that may arise from macrocycle-target binding (Figure 3 and Figure S9). Indeed, in most selections, we observed that those sequences with the fewest sequence counts exhibited both the largest and the smallest enrichment factors (Figure 3 and Figure S9), presumably due to statistical noise among underrepresented library members. The large number of library members enabled a general relationship between sequence abundance and typical enrichment levels to emerge for each selection. Enrichment factors were, therefore, deemed of interest not only on the basis of their absolute value, but also only if they fell above the typical enrichment factor range observed for sequences of that abundance.

Protein Kinase Selection Results. Because most of the significantly enriched DNA sequences arose from selections against protein kinases, we focused our characterization efforts on this class of targets. After one round of selection against Src kinase, macrocycle A11-B1-C5-D7 was enriched 62-fold,

(44) Heid, C. A.; Stevens, J.; Livak, K. J.; Williams, P. M. *Genome Res.* **1996**, *6*, 986–994.
(45) Higuchi, R.; Dollinger, G.; Walsh, P. S.; Griffith, P. R. *Biotechnology* **1992**, *10*, 413–417.
(46) Livak, K. J.; Flood, S. J. A.; Marmaro, J.; Giusti, W.; Deetz, K. *PCR Methods Appl.* **1995**, *4*, 357–362.
(47) Church, G. M. *Sci. Am.* **2006**, *294*, 46–54.
(48) Hall, N. *J. Exp. Biol.* **2007**, *210*, 1518–1525.
(49) Bentley, D. R.; Balasubramanian, S.; Swerdlow, H. P.; Smith, G. P.; Milton, J.; Brown, C. G.; Hall, K. P.; Evers, D. J.; Barnes, C. L.; Bignell, H. R. *Nature* **2008**, *456*, 53–59.

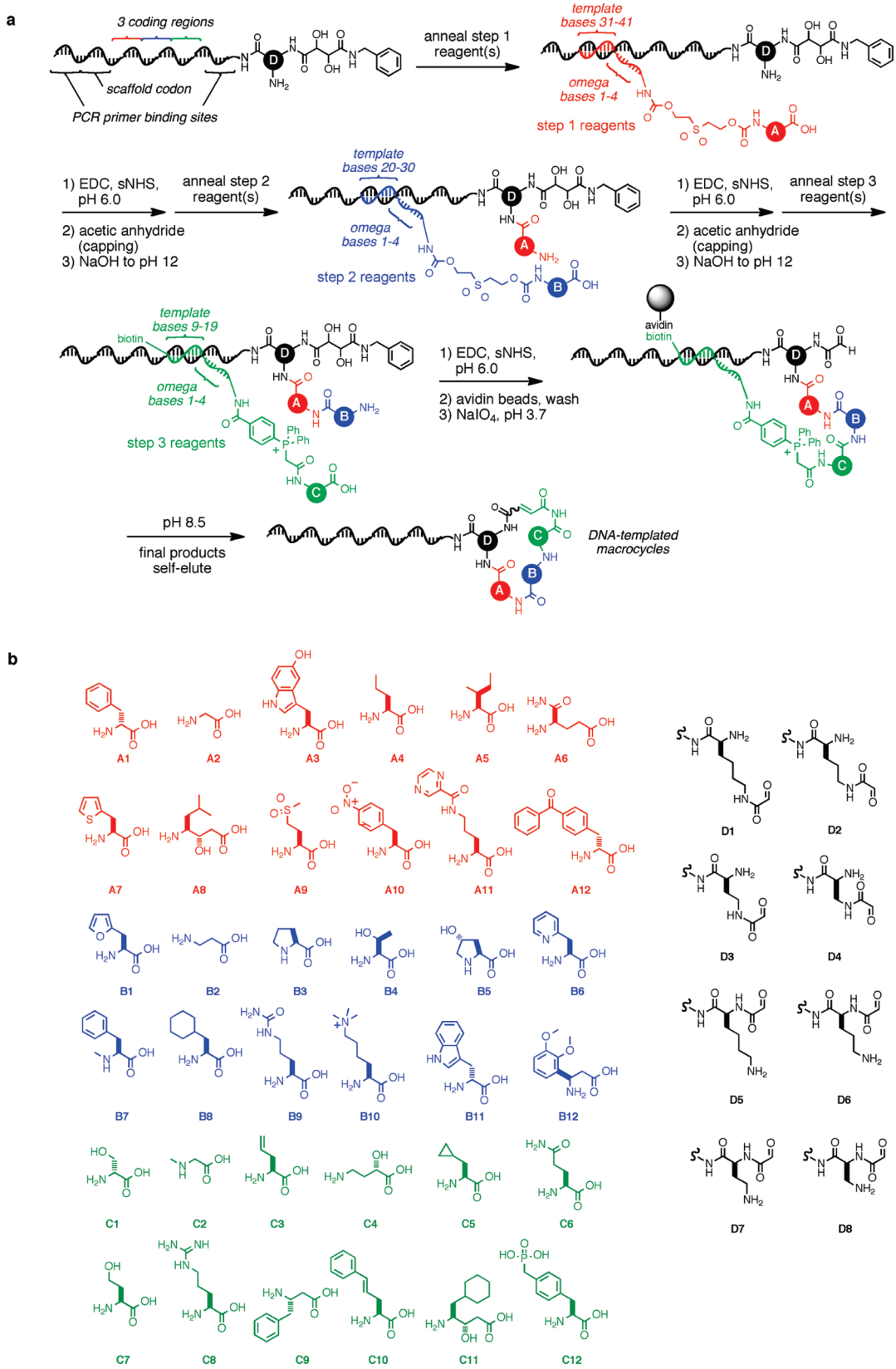


Figure 1. The 13 824-membered DNA-templated small-molecule macrocycle library. (a) Scheme for the multistep DNA-templated synthesis of the macrocycle library.[24] (b) Amino-acid building blocks used in the library synthesis. Step one building blocks (**A1**−**A12**) are shown in red, step two building blocks (**B1**−**B12**) are shown in blue, step three building blocks (**C1**−**C12**) are shown in green, and the macrocycle scaffolds (**D1**−**D8**) are shown in black.

representing the highest degree of enrichment for any library member in any of the selections described above. We also identified macrocycles A11-B1-C10-D7, A11-B1-C3-D7, A10-B1-C5-D6, A10-B1-C5-D7, A11-B1-C5-D5, A9-B1-C5-D7, A11-B8-C5-D7, and A11-B8-C10-D7 (Figures 3 and 4) as enriched between 2- and 7-fold in the Src selection. The

Table 1. Protein Targets for *in Vitro* Selection[a]

kinases	phosphatases	GTPases	PDZ domains	SH2 domains	other proteins
Akt3	DEP1	Cdc42	Dvl1-3	Abl1	Bcl-xL
AMPK	MEG2	H-Ras-V12	Erbin	Abl2	BIR3 (XIAP)
ERBB4	PRL2	RhoA	γ1-syntrophin	P85αN	PPARδ
MK2			Magi1	P85αC	
p38α			MUPP1	P55γC	
MKK6			PAR6B		
Pim1			PSD95		
Src			RGS3		
VEGFR2			SAP97		
			Semcap3		
			Shank3		

[a] All proteins were purchased or purified as GST-tagged fusion proteins (Figure S8).

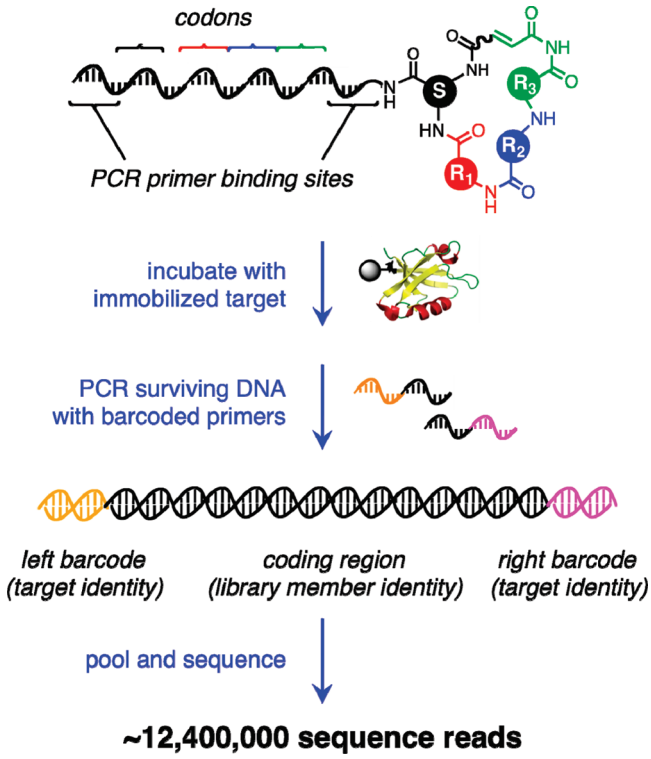


Figure 2. *In vitro* selection of the DNA-templated macrocycle library against 36 protein targets. Following affinity selection of the library against an immobilized GST-fused target protein, library members possessing target affinity were eluted and their attached DNA templates were amplified by PCR using DNA primers containing a barcode sequence that uniquely identifies the protein target. Barcoded PCR amplicons from 36 target protein-binding selections, one control selection for binding immobilized GST, and eight aliquots of unselected library were pooled and submitted as one sample for ultra-high-throughput DNA sequencing.

significant enrichment factors, strong degree of structural similarity among these molecules, and lack of enrichment of these macrocycles in the other eight kinase selections performed (Table 2), collectively suggested that this class of molecules may correspond to authentic Src ligands that are selective for binding to Src over other protein kinases.

We also identified two additional families of macrocycles enriched in selections against several other kinase targets. Macrocycles A10-Y-C11-D5 (where Y = B1 or B8) were enriched against Akt3, MAPKAPK2, Pim1, and VEGFR2 (Figure 5). Macrocycles A12-B8-C10-X (where X = D1−D8) exhibited unusually high enrichment factors in the MAPKAPK2 and Pim1 selections (Figure 5). While we were encouraged by the ability of these molecules to bind similar targets, the high degree of hydrophobicity of the building blocks in these macrocycles raised the possibility that they may bind their targets in a nonclassical mode.[50,51] Moreover, the enrichment of macrocycles of structure A12-B8-C10-X appears to be insensitive to changes in ring size (caused by substitution at the "D" scaffold position) that we anticipated would have large effects on macrocycle conformation, suggesting that these macrocycles were interacting with their targets in a nonclassical binding mode.[24] The possibility that these compounds represent nonspecific ligands or "promiscuous aggregators"[50,51] is explicitly examined below.

Solid-Phase Macrocycle Synthesis. To test if the enriched DNA sequences emerging from selection correspond to macrocycles with target-binding or target-inhibiting activities, we synthesized the corresponding macrocycles on multimilligram scale using Fmoc solid-phase peptide synthesis (Scheme 1).[20] In brief, NovaPEG Rink amide resin (Merck) was coupled to a differentially protected diamino acid building block. An isopropylidene-protected tartrate monomethyl ester was reacted with one amine, and a tripeptide was synthesized on the other amine using three cycles of Fmoc peptide synthesis. After generating the phosphonium salt on resin, we cleaved the macrocyclization precursor from resin using strong acid. This cleavage reaction simultaneously generated a carboxamide at the C-terminus and also revealed the tartrate diol. After HPLC purification and LC/MS characterization of the linear product, we oxidiatively cleaved the diol to reveal an aldehyde and effected Wittig cyclization by raising the pH of the solution to generate a phosphonium ylide. Macrocycle syntheses typically yielded two HPLC-separable isomers that were characterized as the desired *cis* and *trans* macrocycles using NMR spectroscopy and high-resolution mass spectrometry (Table S3). In total, 27 macrocycles corresponding to enriched DNA sequences were synthesized on 1−3 mg scale using this route. Overall yields for each 14-step synthesis (13 solid-phase steps and one solution-phase step) ranged from 1−12%.

Macrocycles Selected for Binding to Src Inhibit Src Kinase Activity. Nine macrocycles enriched in the Src selection (Figure 4) and synthesized as described above were assayed *in vitro* for the ability to inhibit Src kinase activity. Kinase activity was assayed in the presence of varying concentrations of macrocycle using the FRET-based Z′-LYTE assay (Invitrogen). Although our selections did not explicitly select for target inhibition, all but one of the nine macrocycles tested inhibited Src kinase activity (Table 3 and Figure S2). All of the assayed macrocycles with Src selection enrichment factors ≥4-fold exhibited IC_{50} values in the range of ~500 nM to 10 μM. The most potent molecules, macrocycles *trans*-A10-B1-C5-D6 and *cis*-A11-B1-C5-D7, which share the furylalanine (B1) and cyclopropylalanine (C5) building blocks, but otherwise display significant structural diversity, inhibited Src with IC_{50} values of 680 and 960 nM, respectively (Table 3 and Figure S2). We consistently observed large differences in potency between macrocycle *cis* and *trans* stereoisomers (Table 3), highlighting the importance of macrocycle stereochemistry and conformation for activity. For example, the *cis* and *trans* stereoisomers of macrocycles A11-B1-C5-D7 and A10-B1-C5-D6 exhibited a ~10-fold difference in potency (Table 3). Likewise, variants of highly enriched macrocycles such as A11-B1-C5-D7 (62-

(50) Shoichet, B. K. *J. Med. Chem.* **2006**, *49*, 7274–7277.
(51) Coan, K. E.; Shoichet, B. K. *J. Am. Chem. Soc.* **2008**, *130*, 9606–9612.

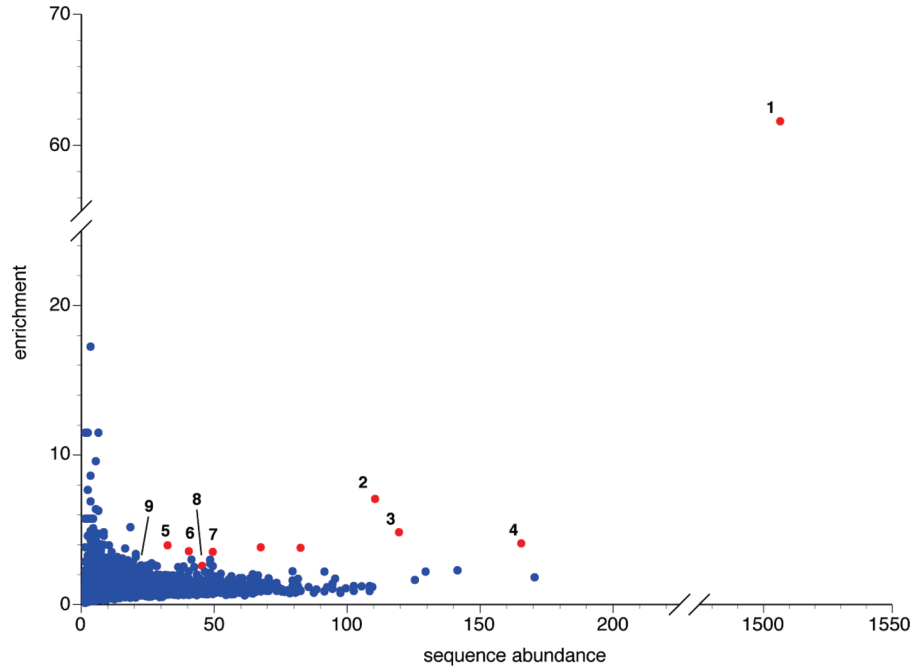


Figure 3. Analysis of high-throughput sequencing results. Plot of enrichment factor vs sequence abundance for library members after selection for binding to Src kinase (enrichment vs sequence abundance plots for all other selections performed are shown in Figure S9). The enrichment factor for each library member in each selection (~497 000 total possible combinations of library member and target) were calculated as (postselection abundance)/(preselection abundance) using a PERL script. Each blue or red dot represents the DNA sequence corresponding to a single library member. The structures of macrocycles **1**−**9** are shown in Figure 4. Enrichment factors for low-abundance library members vary widely due to statistical undersampling. Only enrichment factors that were substantially higher than background values were considered potential positives (shown in red).

fold enrichment) that differ by the addition or loss of a single methylene group within the macrocycle were not significantly enriched, further suggesting that even closely related library members can access significantly different three-dimensional structures. Taken together, these results demonstrate the ability of small molecules emerging from *in vitro* selection of DNA-templated libraries to inhibit with significant potency the protein targets against which they were selected in a manner that is dependent not only on macrocycle building block identity, but also on macrocycle stereochemistry.

Macrocycles Selected for Binding to Akt3, MAPKAPK2, Pim1, and VEGFR2 Modulate Kinase Activity. Macrocycles A10-B1-C11-D5 and A10-B8-C11-D5 were enriched 2- to 6-fold upon selection for binding Akt3, MAPKAPK2, Pim1, and VEGFR2 (Figure 5). These two macrocycles are structurally similar, sharing nitrophenylalanine (A10) and cyclohexylstatine (C11) building blocks as well as the same lysine scaffold (D5), and differ only at the B building block with either furylalanine (B1) or cyclohexylalanine (B8) (Figure 5). For assay purposes, the macrocycles were synthesized with C-terminal PEG-diamine linkers to aid solubility, and then tested for the ability to inhibit Akt3, MAPKAPK2, Pim1 and VEGFR2. The more hydrophobic of the two compounds, macrocycle *cis*-A10-B8-C11-D5, inhibited Akt3, MAPKAPK2, Pim1, and p38α-MAPKAPK2 with IC_{50} values of 8.7, 6.8, 7.5, and 3.1 μM, respectively (Table 4 and Figure S3). Macrocycle *cis*-A10-B1-C11-D5 did not inhibit any of these kinases when assayed with isolated kinase enzymes; however, we observed inhibition of the p38α-MAPKAPK2 kinase cascade with an IC_{50} value of 6.4 μM (Table 4 and Figure S3), suggesting that the molecule may interact with either or both kinases but only in a conformation that is adopted upon formation of the p38α-MAPKAPK2 complex. Interestingly, we observed strong, dose-dependent *activation* of VEGFR2 in the presence of *cis*-A10-B1-C11-D5 (Figure S4) with enzyme activity increasing 70% upon treatment with 10 μM of this macrocycle and 300% upon treatment with 100 μM compound, suggesting that this macrocycle may bind to VEGFR2 in a manner that enhances kinase activity.

We also synthesized and assayed three macrocycles from the A12-B8-C10-X series that were enriched 2- to 10-fold after *in vitro* selection against MAPKAPK2 and Pim1 (Figure 5). These macrocycles possess three hydrophobic building blocks, benzoyl-D-phenylalanine (A12), cyclohexylalanine (B8), and styrylalanine (C10), with variable scaffold diamino acid building blocks. Macrocycles A12-B8-C10-D3, enriched 3-fold against MAPKAPK2 and 3.5-fold against Pim1; A12-B8-C10-D4, enriched 3-fold against MAPKAPK2 and 10-fold against Pim1; and A12-B8-C10-D8, enriched 2.5-fold against MAPKAPK2 and 6-fold against Pim1 (Figure 5), were synthesized and assayed for inhibition of MAPKAPK2, Pim1, and the p38α-MAPKAPK2 cascade. We observed modest inhibition of these kinase targets, with the most potent molecules being *cis*-A12-B8-C10-D3 against the p38α-MAPKAPK2 cascade (IC_{50} = 11 μM) and *cis*-A12-B8-C10-D8 against Pim1 (IC_{50} = 19 μM) (Figure S5 and Table 4).

Taken together, these results indicate that DNA-templated macrocycles that are enriched from *in vitro* selections for kinase affinity frequently possess the ability to inhibit protein kinases with reasonable potency (low- to mid-micromolar IC_{50} values) when synthesized and assayed in a non-DNA-linked form. Among the nine macrocycles chosen based on the magnitude of their enrichment during Src selection, or based on their structurally similarity to the highly enriched A11-B1-C5-D7, we characterized only one enriched macrocycle (A9-B1-C5-D7) that did not inhibit Src, representing a false positive rate of 11%. While we characterized only five molecules from selections against Akt3, MAPKAPK2, Pim1, and VEGFR2, due to enrichment of the same molecule against multiple targets,

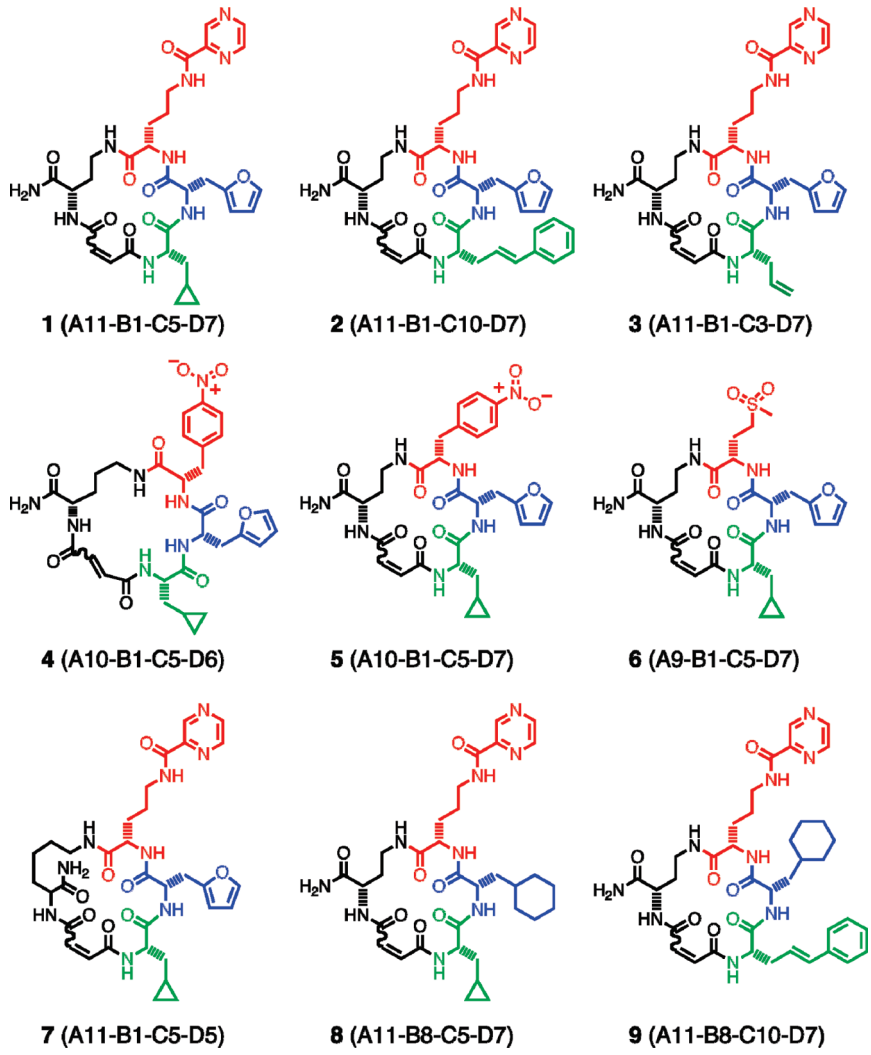


Figure 4. Chemical structures of macrocycles exhibiting significant enrichment above background and common structural motifs after selection against Src kinase. Macrocycle numbering corresponds to that used in Figure 3.

Table 2. Enrichment Factors of Positives from the Src Selection in All Protein Kinase Selections Performed[a]

	Src	Akt3	AMPK	ERBB4	MK2	p38α	MKK6	Pim1	VEGFR2
1	62	2.2	0.69	0.63	2.8	1.1	2.7	2.4	2.4
2	7.1	1.2	0.86	0.82	1.8	0.84	1.1	0.91	0.95
3	4.8	0.97	0.94	0.91	1.2	1.0	0.81	0.93	1.4
4	4.1	0.81	0.8	1.1	0.97	0.91	1.1	0.89	0.65
5	4.0	1.5	1.4	1.7	2.0	1.7	2.2	1.3	1.6
6	3.6	1.3	1.0	1.0	0.65	1.1	1.4	1.1	0.66
7	3.5	1.5	1.1	1.2	1.5	0.99	1.0	0.39	1.4
8	2.6	1.5	1.2	1.2	2.0	1.6	1.3	2.2	1.5
9	2.6	0.98	0.88	1.3	2.6	1.4	2.4	1.9	1.6

[a] Macrocycle numbering is from Figure 4. Enrichment factors >10 are colored red; those >4 are colored orange; those >2 are colored yellow.

we validated 12 distinct protein-macrocycle interactions. One molecule, macrocycle A10-B9-C11-D5 (B9 = citrulline), similar to the validated A10-Y-C11-D5 family of hits, was observed enriched above background in the Pim1 (3.5-fold enrichment) and VEGFR2 (2.4-fold enrichment) selections, but did not inhibit either kinase when assayed in its non-DNA-linked form. In total, we observed a similarly low false positive rate of 14% (2/14) from the Akt3, MAPKAPK2, Pim1, and VEGFR2 selections.

Kinase Inhibition Selectivity of Enriched Src-Inhibiting Macrocycles. We chose to characterize in greater depth macrocycles *cis*-A11-B1-C5-D7 and *trans*-A10-B1-C5-D6 (Figure 4 and Table 3) since they represent potent and structurally diverse inhibitors that are predicted to be selective for Src kinase based on *in vitro* selection results (Table 2). We first characterized their mode of inhibition. Although the selection used to identify the inhibitors is agnostic with respect to target binding

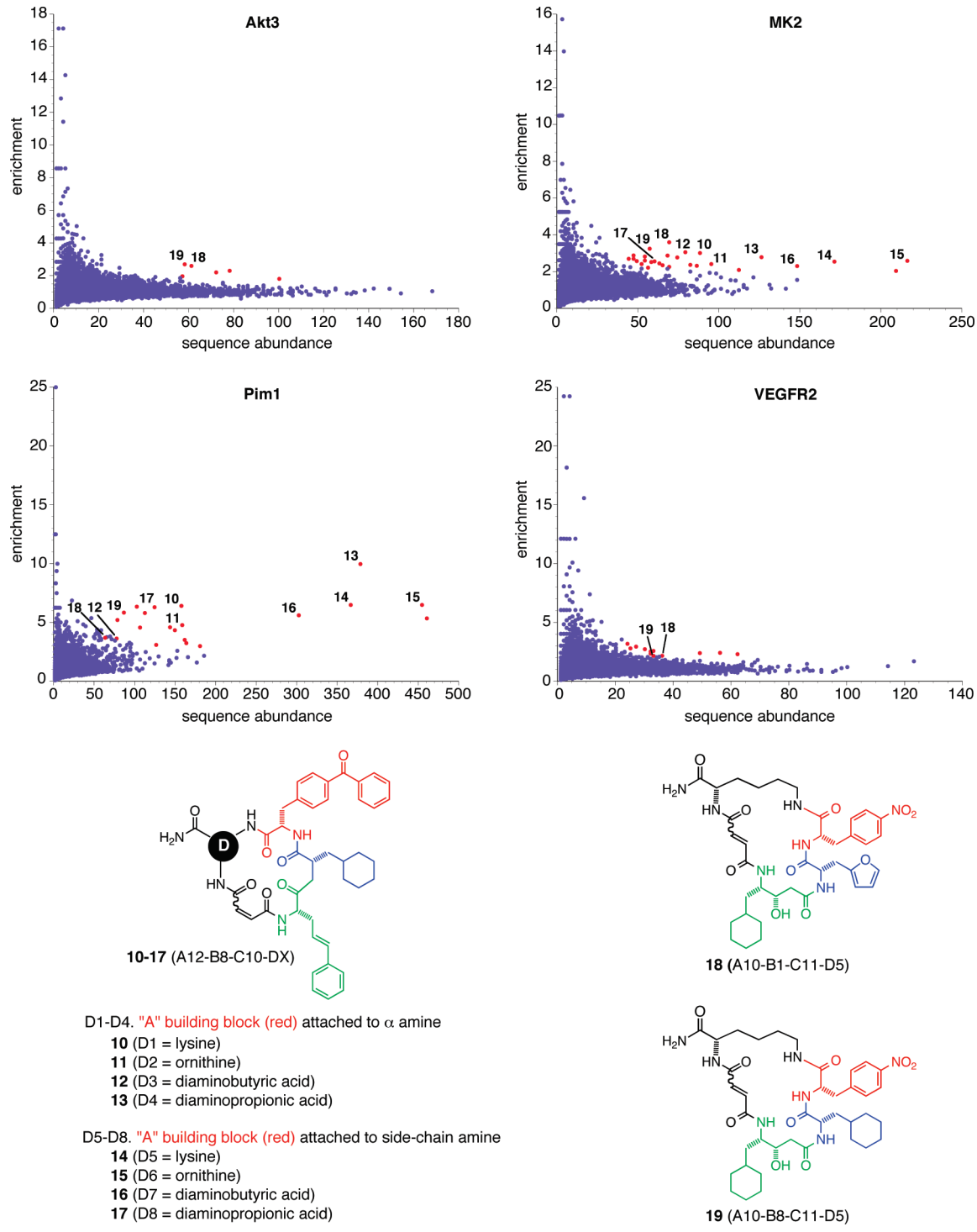


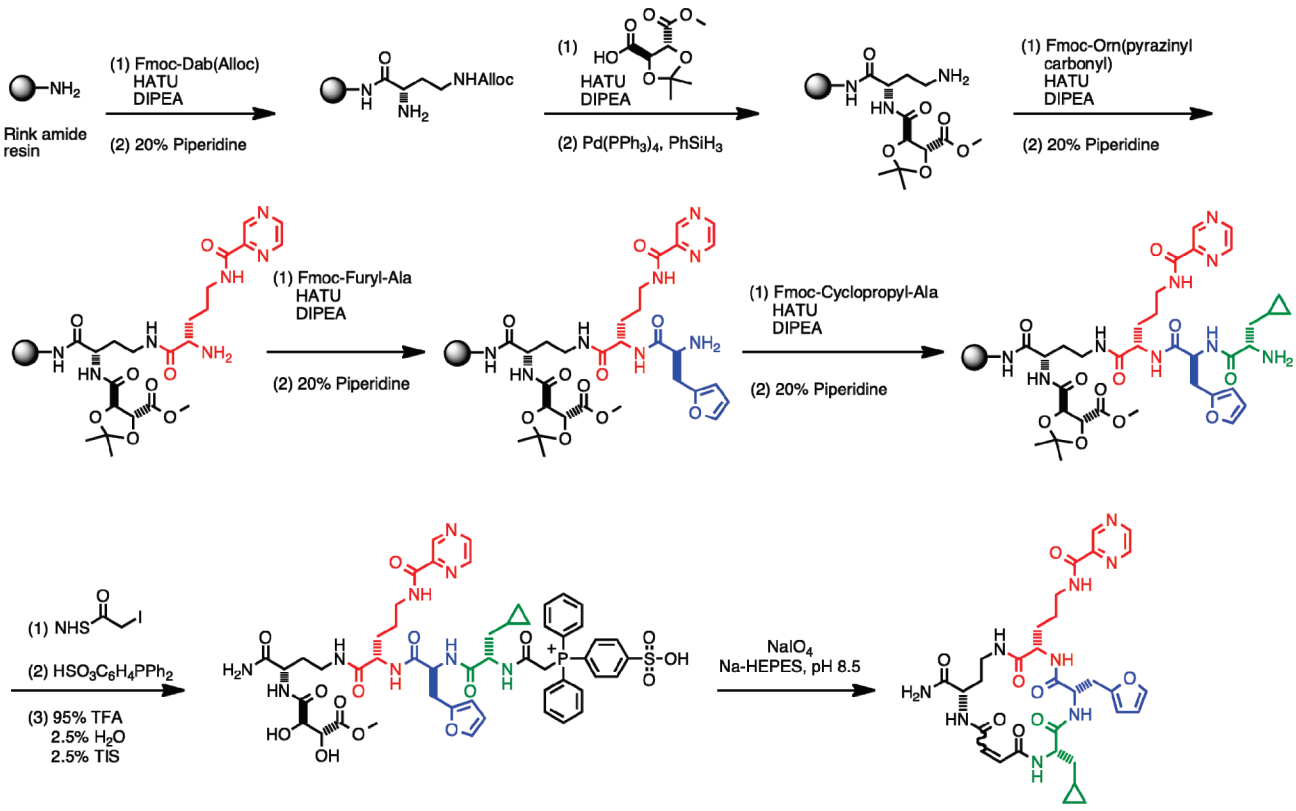
Figure 5. Plots of enrichment factor vs sequence abundance for library members after selection for binding to Akt3, MAPKAPK2, Pim1, and VEGFR2. Macrocycles **10**−**19** exhibit significant enrichment above background.

site, we determined that *cis*-A11-B1-C5-D7 and *trans*-A10-B1-C5-D6 are ATP-competitive inhibitors (Figure 6 and Figure S6). We next determined the selectivity of these molecules against a representative set of human kinases. Although kinase selectivity can be modest for ATP-competitive inhibitors since the ATP-binding site is highly conserved among protein kinases,[52−55] our selection data (Table 2) suggested that the Src-inhibiting macrocycles may exhibit unusual selectivity.

Macrocycle *cis*-A11-B1-C5-D7 was assayed at 50 μM against a representative panel of 58 commercially available human protein kinases including all nine mammalian Src family kinases (Table S4). For those kinases that displayed significant inhibition (>40%) at 50 μM *cis*-A11-B1-C5-D7, we reassayed at 5 μM compound concentration to estimate rough IC_{50} values and performed 10-point IC_{50} titrations for the most potently inhibited

(52) Hanks, S. K.; Hunter, T. *FASEB J.* **1995**, *9*, 576–596.
(53) Hanks, S. K.; Quinn, A. M. *Methods Enzymol.* **1991**, *200*, 38–62.
(54) Hanks, S. K.; Quinn, A. M.; Hunter, T. *Science* **1988**, *241*, 42–52.
(55) Stout, T. J.; Foster, P. G.; Mathews, D. J. *Curr. Pharm. Des.* **2004**, *10*, 1069–1082.

Scheme 1. Milligram-Scale Synthesis of Macrocycle A11-B1-C5-D7[a]



[a] The route was adapted from Gartner et al.[20] Methyl hydrogen 2,3-*O*-isopropylidene-(L)-tartrate was synthesized as reported by Musich and Rapoport.[70]

Table 3. Macrocycles Corresponding to Src Selection Survivors Inhibit Src Kinase[a]

macrocycle	Src IC_{50} (μM)
cis-A11-B1-C5-D7	0.96
trans-A11-B1-C5-D7	12
cis-A11-B1-C10-D7	7.0
trans-A11-B1-C10-D7	21
cis-A11-B1-C3-D7	10
trans-A11-B1-C3-D7	48
cis-A10-B1-C5-D6	7.4
trans-A10-B1-C5-D6	0.68
cis-A10-B1-C5-D7	8.9
trans-A10-B1-C5-D7	1.5
cis-A9-B1-C5-D7	>100
trans-A9-B1-C5-D7	>100
cis-A11-B1-C5-D5	85
trans-A11-B1-C5-D5	>100
cis-A11-B8-C5-D7	36
trans-A11-B8-C5-D7	>100
cis-A11-B8-C10-D7	25
trans-A11-B8-C10-D7	>100

[a] Src kinase activity was assayed in the presence of increasing concentrations of macrocycle using the Z′-LYTE assay (Invitrogen).

kinases (Table S4). Using this data, we calculated a selectivity score[56] for the macrocycle based on the number of nontarget proteins exhibiting IC_{50} values within 10-fold of the 960 nM IC_{50} of *cis*-A11-B1-C5-D7 for Src. Macrocycle *cis*-A11-B1-C5-D7 displayed a selectivity score of 0.05, indicating that only 5% of the kinases assayed were inhibited with IC_{50} values within

Table 4. Enriched Macrocycles Inhibit Pim1, MK2, Akt3, and p38a-MK2[a]

macrocycle	kinase	IC_{50} (μM)
cis-A10-B1-C11-D5	p38α-MK2	6.5
cis-A10-B1-C11-D5	Pim1	21
cis-A10-B8-C11-D5	Akt3	8.4
cis-A10-B8-C11-D5	MK2	6.1
cis-A10-B8-C11-D5	p38α-MK2	3.4
cis-A10-B8-C11-D5	Pim1	7.3
cis-A12-B8-C10-D3	p38α-MK2	12
cis-A12-B8-C10-D3	Pim1	25
cis-A12-B8-C10-D4	p38α-MK2	20
cis-A12-B8-C10-D4	Pim1	100
cis-A12-B8-C10-D8	MK2	47
cis-A12-B8-C10-D8	p38α-MK2	14
cis-A12-B8-C10-D8	Pim1	17

[a] Kinase activity was assayed in the presence of increasing concentrations of macrocycles using Invitrogen's Z-LYTE assay.

10-fold of the IC_{50} for Src. When judged by similar criteria, *cis*-A11-B1-C5-D7 is significantly more selective than the promiscuous ATP-competitive inhibitor staurosporine (selectivity score = 0.50), and slightly more selective than two kinase inhibitor drugs: sorafenib (selectivity score = 0.10) and dasatanib (selectivity score = 0.10).[56] The most potently inhibited proteins other than Src were closely related kinases such as Fgr (a Src family kinase, IC_{50} = 4.0 μM) and RET (a tyrosine kinase, IC_{50} = 5.7 μM), as well as Aurora A, a serine/threonine kinase which shares close structural homology to Src-family kinases (IC_{50} = 1.7 μM).[57]

(56) Karaman, M. W.; Herrgard, S.; Treiber, D. K.; Gallant, P.; Atteridge, C. E.; Campbell, B. T.; Chan, K. W.; Ciceri, P.; Davis, M. I.; Edeen, P.; Faraoni, R.; Floyd, M.; Hunt, J. P.; Lockhart, D. J.; Milanov, Z. V.; Morrison, M. J.; Pallares, G.; Patel, H. K.; Pritchard, S.; Wodicka, L. M.; Zarrinkar, P. P. *Nat. Biotech.* **2008**, *26*, 127–132.

(57) Cheetham, G. M. T.; Knegtel, R. M. A.; Coll, J. T.; Renwick, S. B.; Swenson, L.; Weber, P.; Lippke, J. A.; Austen, D. A. *J. Biol. Chem.* **2002**, *277*, 42419–42422.

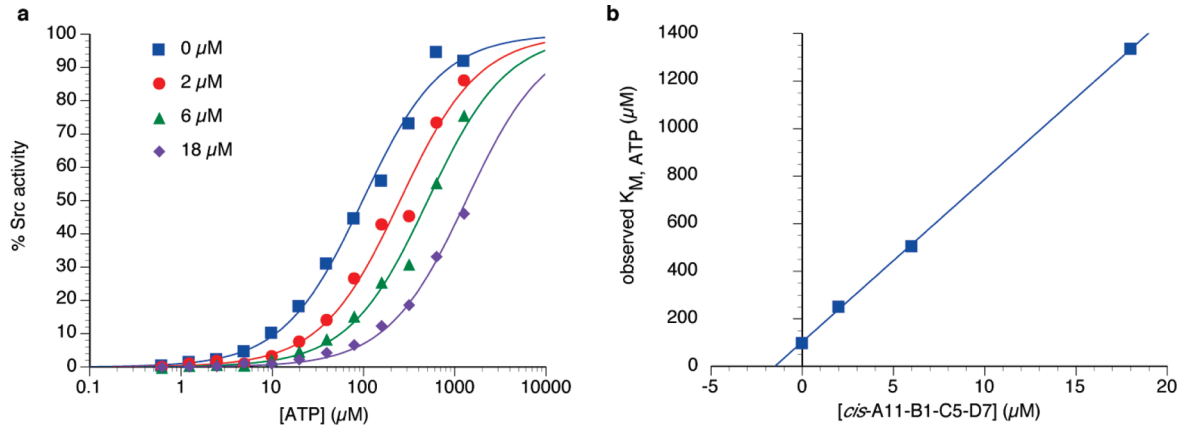


Figure 6. Macrocycle *cis*-A11-B1-C5-D7 is an ATP-competitive inhibitor of Src kinase. (a) The apparent K_M of Src for ATP was measured in the presence of increasing concentrations of macrocycle *cis*-A11-B1-C5-D7. Kinase activity was measured as in Table 3. (b) The resulting relationship is linear in accord with the following equation for a classical competitive inhibitor: apparent $K_M = K_M \times (1 + [\text{inhibitor}]/K_i)$.

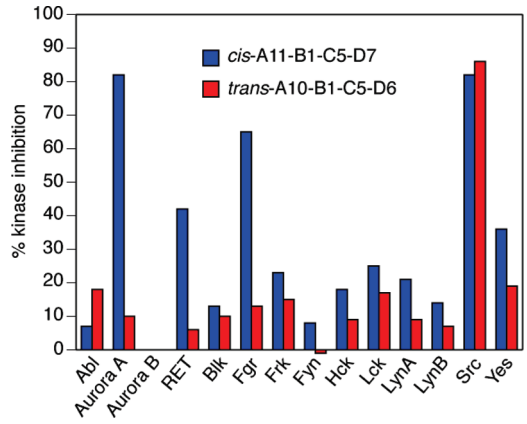


Figure 7. Selectivity of macrocycles *cis*-A11-B1-C5-D7 and *trans*-A10-B1-C5-D6 among Src-related kinases, including all nine Src-family kinases. Both macrocycles were assayed at 5 μM concentration against the indicated kinases. Assays of *cis*-A11-B1-C5-D7 against Abl and Hck were performed at 50 μM macrocycle concentration instead of 5 μM. Percent kinase inhibition values plotted are the average of two independent measurements. Assay data for Src kinase is from Figure S2 and Table 3. All non-Src assay points were measured by the Invitrogen Select Screen Profiling Service.

In addition to good overall selectivity, *cis*-A11-B1-C5-D7 exhibited exceptional selectivity among Src-like protein kinases that have traditionally been difficult to distinguish by small-molecule inhibition (Figure 7). For example, this macrocycle exhibits >50-fold selectivity for Src over Abl kinase, > 50-fold selectivity for Src over the Src-family kinases Hck and Fyn, and >10-fold selectivity for Src over all remaining Src-family kinases with the exception of Fgr (Figure 7 and Table S4). Similarly, among the closely related Aurora kinases, we observed >50-fold selectivity for Aurora A over Aurora B (Figure 7 and Table S4).

We also investigated the selectivity of the nitrophenylalanine (A10)-containing macrocycle *trans*-A10-B1-C5-D6. While it shares two out of four building blocks in common with *cis*-A11-B1-C5-D7 and also inhibits Src in an ATP-competitive manner, building block composition in this macrocycle is strikingly different at the A position and the *trans* stereochemistry of the olefin contrasts with the olefin geometry of the active isomer of A11-B1-C5-D7. The differences between these two molecules were further revealed upon assays against a kinase panel, in which *trans*-A10-B1-C5-D6 exhibited even greater selectivity for Src than macrocycle *cis*-A11-B1-C5-D7 (Figure 7). Macrocycle *trans*-A10-B1-C5-D6 did not significantly inhibit Aurora A, RET, or any of the other Src-family kinases (the primary kinases beyond Src that are inhibited by *cis*-A11-B1-C5-D7) at the concentration tested (5 μM) (Figure 7). Indeed, of all 44 human kinases against which this macrocycle was assayed, Src was the only kinase observed to be significantly inhibited by this macrocycle (Figure 7 and Table S5). This level of Src selectivity is unprecedented among currently available Src-selective kinase inhibitors such as PP1 and PP2,[58] which are significantly more potent against Src-family members Fyn, Hck, and Lyn than against Src itself. Taken together, these results establish that macrocycles emerging from the *in vitro* selection of a DNA-templated small-molecule library include kinase inhibitors that exhibit unusual selectivity.

SAR Analysis of Src-Inhibiting Macrocycles. We analyzed the set of macrocycles enriched in the Src selection to develop basic structure−activity relationships for Src inhibition. Our selection data shows a strong preference for a small set of building blocks at each nonscaffold position (Figure 4), suggesting that all three macrocyclic building blocks are important for binding. To test this hypothesis, we synthesized three derivatives of macrocycle *cis*-A11-B1-C5-D7 by systematically replacing each of the nonscaffold building blocks (positions A, B, and C) with alanine, in analogy to alanine-scanning mutagenesis of proteins. We assayed the resulting three single-alanine mutants for Src inhibition and observed that none of the three mutant macrocycles at 5 μM significantly inhibit Src or the Src-family kinase Fgr (Figure 8a), indicating that all three building blocks contribute to the kinase inhibition activity of *cis*-A11-B1-C5-D7.

To evaluate the importance of the macrocyclic nature of these molecules for kinase inhibition, we synthesized two diacetylated linear peptides as linear versions of macrocycles A11-B1-C5-D7 and A10-B1-C5-D6 (Figure 8b) and assayed their abilities to inhibit Src. These two compounds are identical to their corresponding macrocycles with the exception that the olefin group resulting from Wittig macrocyclization is replaced by two methyl groups in the linear peptides. Both linear peptides were virtually inactive against Src kinase, confirming the necessity of a cyclic scaffold to preorganize these sets of building blocks

(58) Hanke, J. H.; Gardner, J. P.; Dow, R. L.; Changelian, P. S.; Brissette, W. H.; Weringer, E. J.; Pollok, B. A.; Connelly, P. A. *J. Biol. Chem.* **1996**, *271*, 695–701.

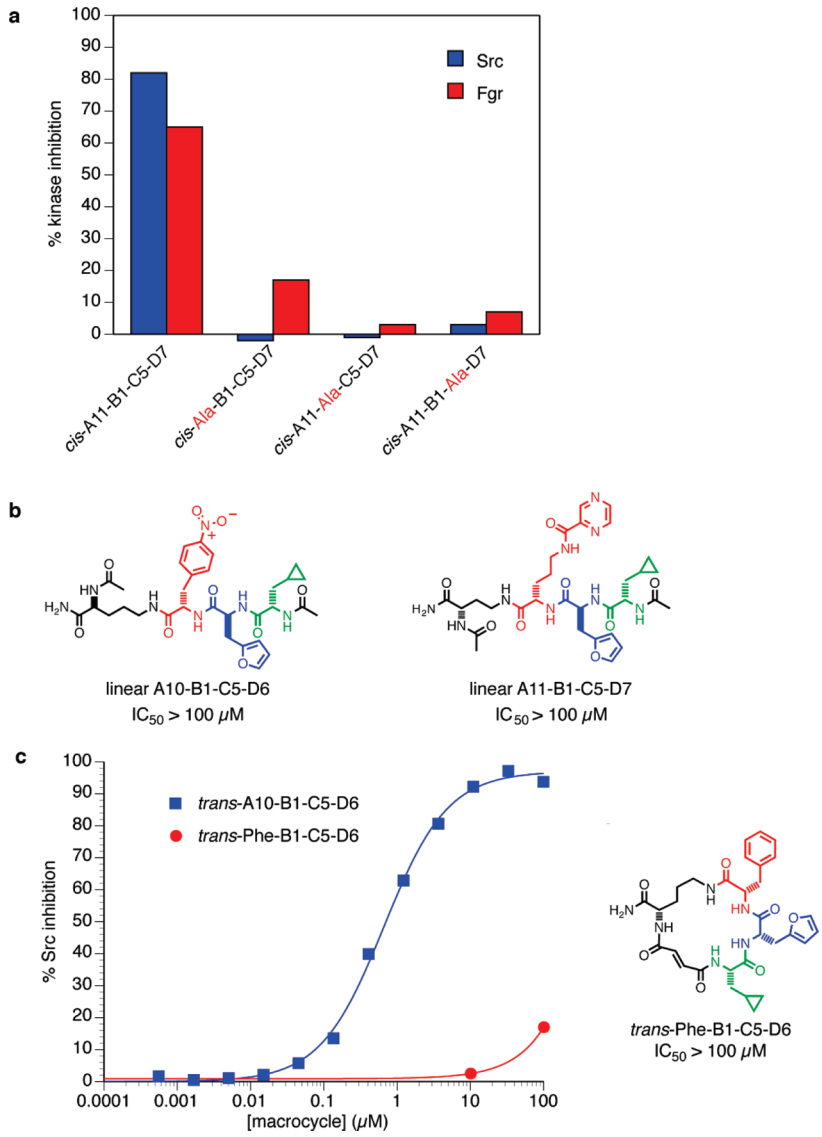


Figure 8. SAR analysis of Src-inhibiting macrocycles. (a) Single-alanine mutants of *cis*-A11-B1-C5-D7 were assayed against Src-family kinases Src and Fgr at 5 μM concentration as described previously. (b) Linear diacetylated peptides corresponding to macrocycles A11-B1-C5-D7 and A10-B1-C5-D6 were synthesized by Fmoc solid-phase peptide synthesis and assayed against Src kinase. (c) Macrocycle *trans*-Phe-B1-C5-D6 (replacing nitrophenylalanine with phenylalanine) was assayed against Src as described previously.

into a three-dimensional conformation capable of binding to and inhibiting the kinase.

Olefin stereochemistry plays a critical role in the ability of these compounds to inhibit Src kinase activity. Macrocycles possessing the ornithine(pyrazinylcarbonyl) building block (A11) and diaminobutyric acid scaffold (D7) are more potent Src inhibitors in their *cis*-isomeric form (Table 3). These macrocycles favor furylalanine (B1) over cyclohexylalanine (B8) at the B position, as evidenced by the weakened potency of macrocycles *cis*-A11-B8-C5-D7 and *cis*-A11-B8-C10-D7 compared to *cis*-A11-B1-C5-D7 and *cis*-A11-B1-C10-D7 (Table 3), respectively. All of these compounds exhibit hydrophobic building blocks at the C position, especially cyclopropylalanine (C5) (Figure 4 and Table 3).

In contrast, macrocycles with nitrophenylalanine (A10) at the A position are more active Src inhibitors in their *trans*-isomeric form (Table 3). The ornithine (D6) scaffold also supports more potent Src inhibition than the diaminobutyric acid (D7) scaffold that differs by the loss of a single methylene group (Table 3). We assayed the importance of the nitrophenylalanine (A10) building block for Src inhibition by replacing it with phenylalanine in macrocycle *trans*-A10-B1-C5-D6 to create macrocycle *trans*-Phe-B1-C5-D6. This substitution abolished Src inhibition (Figure 8c), indicating the importance of the nitro group among A10 macrocycle subfamily kinase inhibitors. Because of the altered olefin stereochemistry of the nitrophenylalanine (A10)-containing macrocycles, the difference in macrocycle ring size, and the structural differences between the nitrophenylalanine (A10) and ornithine(pyrazinylcarbonyl) (A11) building blocks, we believe that these two classes of macrocycles interact in distinct ways with the Src active site. Indeed, modeling the three-dimensional conformations adopted by *cis*-A11-B1-C5-D7 and *trans*-A10-B1-C5-D6 using an MM2 energy-minimization algorithm (Chem3D, CambridgeSoft) predicts large differences in both preferred backbone conformation and preferred side-chain orientation for these two macrocycles (Figure S10). This hypothesis is further supported by the substantial differences in kinase selectivity (Figure 7) exhibited by *cis*-A11-B1-C5-D7 and *trans*-A10-B1-C5-D6.

Src-Inhibiting Macrocycles Do Not Resemble Promiscuous Aggregators. High-throughput screening can result in the discovery of noncanonical ligands known as promiscuous aggregators that form colloidal aggregates in aqueous solution and inhibit enzymes by a nonspecific sequestration mechanism rather than through a classical one-to-one binding mode.[50] A common property of promiscuous aggregators is a steep dose−response curve that can result from the enzyme concentration in the kinase assay exceeding the K_i of the aggregated species, which is typically present at significantly lower concentrations than that of the small molecule itself.[50] We suspected that A10-Y-C11-D5 and A12-B8-C10-X macrocycles (Figure 5 and Table 4) might be promiscuous aggregators due to their high hydrophobicity, broad specificity, and steep dose−response curves (Figures S3 and S5).

To test this hypothesis, we assayed macrocycle *cis*-A10-B8-C11-D5 against Akt3, MAPKAPK2, and Pim1 and assayed macrocycle *cis*-A12-B8-C10-D3 against Pim1 while varying the concentration of kinase in the reactions. If aggregate formation results in stoichiometric titration of the kinase, the observed IC_{50} value should vary linearly with enzyme concentration.[50] Despite their steep dose−response curves, inhibition of MAPKAPK2 and Pim1 did not vary with kinase concentration (Figure S7), suggesting a different mechanism of inhibition. For example, steep dose−response curves can also be caused by multisite binding or by a phase transition in the inhibitor that is coupled with inhibition.[50] In contrast, we observed that inhibition of Akt3 by *cis*-A10-B8-C11-D5 varied linearly with the kinase concentration used in the assay (Figure S7), indicating that this macrocycle may be inhibiting Akt3 through a promiscuous aggregator mechanism.

In contrast, the family of Src-inhibiting macrocycles discovered in this work do not exhibit steep dose−response curves (Figure S2), bind in an ATP-competitive manner (Figure 6 and Figure S6), and do not vary significantly in observed inhibitory potency as a function of enzyme concentration (Figure S7). These results collectively indicate that the Src-inhibiting macrocycles described above are not promiscuous aggregators.

Discussion

DNA-encoded small-molecule libraries can be subjected to simple affinity-based selection using disposable tubes, microgram quantities of immobilized protein targets, and micropipets, in contrast with the more labor- and infrastructure-intensive process of conventional high-throughput screening. As a result, as demonstrated in this work, DNA-encoded small-molecule ligands for proteins of interest can be revealed in a manner that is fast, inexpensive, and largely independent of library size. Here, we subjected a 13 824-membered DNA-templated macrocycle library to 36 protein affinity-based selections in a parallel manner. These selections collectively queried ~497 000 protein-small molecule interactions in a process that required 3 days of one researcher's effort to complete.

The parallel, reasonably low-cost (in 2009, ~$5500) analysis of the 36 selections was made possible by recent developments in ultra-high-throughput sequencing technology that provided millions of sequence reads of sufficient length to span the coding regions of our library templates. PCR-installed DNA barcodes enabled all of the selections to be sequenced as one sample and deconvoluted for analysis by binning sequence data by barcode. Our results validate the ability of current DNA sequencing technology to detect even modestly enriched library members, obviating the need to perform multiple rounds of selection for the library members characterized in this work. The low false positive rate among the macrocycles enriched during selections studied here suggests that the evaluation of small molecule−protein interactions based on *in vitro* selection and modern DNA sequencing methods is robust. As ultra-high-throughput DNA sequencing continues to decrease rapidly in cost and increase in capability, the analysis of *in vitro* selections evaluating millions of potential DNA-encoded small molecule−target combinations will become increasingly efficient and routine.

After only one round of *in vitro* binding selection, high-throughput sequencing enabled the identification of three families of enriched macrocycles in the protein kinase selections, despite enrichment factors that were as low as ~2- to 3-fold. All three families of macrocycles enriched after selection were shown to possess the ability to inhibit their target kinase, although we chose not to further characterize two of the three macrocycle families due to their modest potencies, promiscuity, or poor physicochemical properties. In contrast, macrocycles enriched in the Src selection (which also displayed the largest enrichment factors of any of our selections) possessed classical single-site dose response curves with IC_{50} values as low as 680 nM. The potency of these compounds is promising given that they emerged from the broad selection of an untargeted 13 824-membered library.

Consistent with the highly target-specific *in vitro* selection results, assays of Src-enriched macrocycles against a broad set of human kinases revealed that these molecules indeed possess very good overall selectivity among a panel of human kinases. The specificity of these macrocycles is especially intriguing since these macrocycles are ATP-competitive inhibitors and the ATP-binding site is highly conserved among protein kinases.[52−54] Furthermore, in addition to possessing good overall selectivity, the molecules were exceptionally selective among kinases that have been traditionally very difficult to distinguish by small-molecule inhibitors. For example, Src and Abl are two closely related nonreceptor tyrosine kinases that have exhibited very similar small-molecule binding specificities.[59−62] Toward the goal of distinguishing Src and Abl using a small molecule, Maly and co-workers[63] recently generated bivalent kinase inhibitors that rely on peptidic recognition outside of the ATP-binding cleft to generate binding selectivity. Macrocycle A11-B1-C5-D7 displays >50-fold selectivity for Src over Abl kinase while binding in an ATP-competitive manner. We also observed >100-fold selectivity for Aurora A among the closely related Aurora family of serine/threonine protein kinases.[64] Most reported small molecule Aurora A inhibitors also inhibit Aurora B and/or Aurora C with comparable affinity,[65−68] although examples of Aurora A selective inhibitors have begun to emerge.[69]

Macrocycle *trans*-A10-B1-C5-D6 exhibited even better selectivity for Src, with no significant off-target inhibition in the panel of 44 human kinases against which this compound was

(59) Azam, M.; Nardi, V.; Shakespeare, W. C.; Metcalf, C. A., III; Bohacek, R. S.; Wang, Y.; Sundaramoorthi, R.; Sliz, P.; Veach, D. R.; Bornmann, W. G.; Clarkson, B.; Dalgarno, D. C.; Sawyer, T. K.; Daley, G. Q. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9244–9249.
(60) Das, J.; et al. *J. Med. Chem.* **2006**, *49*, 6819–6832.
(61) Golas, J. M.; Lucas, J.; Etienne, C.; Golas, J.; Discafani, C.; Sridharan, L.; Boghaert, E.; Arndt, K.; Ye, F.; Boschelli, D. H.; Li, F.; Titsch, C.; Huselton, C.; Chaudary, I.; Boschelli, F. *Cancer Res.* **2005**, *65*, 5358–5364.
(62) Tatton, L.; Morley, G. M.; Chopra, R.; Khwaja, A. *J. Biol. Chem.* **2003**, *278*, 4847–4853.
(63) Hill, Z. B.; Gayani, B.; Perara, K.; Maly, D. J. *J. Am. Chem. Soc.* **2009**, *131*, 6686–6688.
(64) Fu, J.; Bian, M.; Jian, Q.; Zhang, C. *Mol. Cancer Res.* **2007**, *5*, 1–10.

assayed. This panel included all nine highly related Src-family kinases. The generation of small molecules that are selective among the Src-family kinases has posed a significant challenge to academic and pharmaceutical drug-discovery efforts, and the new class of macrocyclic kinase inhibitors described here represent promising starting points for future probe or therapeutic development efforts.

It is tempting to speculate that the observed selectivity arises at least in part from the design of the library to contain rigid macrocyclic structures with a high ratio of atoms derived from varying building blocks to atoms shared in common by all library members. Kinase selectivity may also result from the larger size of these molecules compared with traditional ATP-competitive kinase inhibitors; the presence of additional groups may support interactions both inside and outside of the ATP-binding site. The macrocyclic kinase inhibitors discovered in this work share very few chemical features with classical ATP-competitive kinase inhibitors, which are typically planar heterocycles based on adenine. In contrast, our peptidic macrocycles contain several stereocenters and can adopt distinct three-dimensional conformations required for selective target binding. These findings also demonstrate that changes in building block composition that do not abrogate Src inhibition can modulate the selectivity of these macrocycles among human kinases.

Conclusion

We have completed a large-scale *in vitro* selection effort of a DNA-templated small-molecule macrocycle library against a panel of 36 diverse protein targets. The selections yielded a new molecular scaffold for Src kinase inhibition that exhibits well-defined structure−activity relationships and exceptional selectivity among human kinases. These findings validate the use of DNA-templated libraries integrated with *in vitro* selection for the discovery of synthetic small molecules with the ability to potently and selectively modulate the activity of protein targets, and also reveal promising compounds for the further development of selective and potent molecules that can modulate the function of cellular kinases.

Materials and Methods

***In Vitro* Selections with Known Protein-Binding Ligands.** A stock solution of GST-labeled protein target was diluted in protein-binding buffer (50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 10 mM β-mercaptoethanol) to a final concentration of 1 μM. The protein target (200 pmol) was incubated with 2 μL of MagneGST (Promega) glutathione-linked magnetic particles for 1 h at 4 °C with gentle shaking. The beads were washed three times with 100 μL of TBST buffer supplemented with 1 M NaCl, followed by two washes with 300 μL of protein-binding buffer. The beads were then diluted in 20 μL selection buffer (TBST, 3 mg/mL yeast tRNA, 1 mM DTT) and 10 μL of resulting protein bead solution was combined with 1 pmol of a nonbinding DNA sequence and 1/10, 1/100, 1/1000, or 1/10 000 pmol of the known protein-binding ligand conjugated to DNA. After removing unbound molecules, the beads were washed three times with 200 μL of selection buffer, and bound molecules were eluted in 20 μL of 0.1 mM glutathione in 50 mM Tris pH 8.0 buffer for 15 min. Of this eluant, 5 μL was subjected to PCR amplification.

Positive Control Selection Analysis. Positive-control selections were analyzed using either a *Taqman* qPCR assay or restriction-endonuclease digestion. *Taqman* qPCR was performed according to the manufacturer's protocols. For restriction-endonuclease analysis, 20 μL of PCR reaction were combined with 5 units of *Ear*I restriction endonuclease (NEB), and incubated for 2 h at 37 °C. Digested PCR amplicons were separated by agarose gel electrophoresis and stained with EtBr. The ratio of digested to undigested DNA was determined by ethidium bromide staining and densitometry using UV light.

***In Vitro* Selections with the DNA-Templated Macrocycle Library.** *In vitro* selections of the 13 824-membered library were conducted exactly as described above for the positive control selections. For each selection, 5 pmol of library was used. PCR amplification of selection survivors was performed with 5′-barcoded forward and reverse primers. Barcoded PCR amplicons were quantitated using Picogreen dsDNA quantitation reagent (Invitrogen), mixed in equimolar amounts, and submitted for high-throughput DNA sequencing.

Kinase Assays. Src IC_{50} determinations and single-point inhibition measurements were performed using Invitrogen's Z′-LYTE Tyr 02 assay kit according to the manufacturer's protocols. All kinase assays were performed at ATP concentrations near $K_{M,ATP}$ values, except where noted otherwise. Active human Src kinase was obtained from Millipore. For kinome profiling and IC_{50} determinations against non-Src kinases, compounds were submitted to Invitrogen's SelectScreen Kinase Profiling Service.

Acknowledgment. This research was supported by the NIH and NIGMS (R01 GM065865) and the Howard Hughes Medical Institute. We thank Gavin MacBeath, Michael Stiffler and Jack Allen for kindly providing purified GST-tagged PDZ and SH2 domains. We thank Carsten Russ and Heidi Spurling for providing high-throughout DNA sequencing, and we thank Walter Kowtoniuk for assistance with mass spectrometry. R.E.K. gratefully acknowledges NIH training grant support to the Harvard University Training Program in Molecular, Cellular, and Chemical Biology (MCCB). C.E.D. gratefully acknowledges the support of the Novartis Foundation and the Swiss National Science Foundation.

Supporting Information Available: NMR spectra, LC/MS data, positive-control selections data, enrichment profiles of library selections, macrocycle titrations, ATP titrations, complete human kinase assay data, and complete refs 39, 60, 67 and 69. This material is available free of charge via the Internet at http://pubs.acs.org.

JA104903X

(65) Andersen, C. B.; Wan, Y.; Chang, J. W.; Riggs, B.; Lee, C.; Liu, Y.; Sessa, F.; Villa, F.; Kwiatkowski, N.; Suzuki, M.; Nallan, L.; Heald, R.; Musacchio, A.; Gray, N. S. *ACS Chem. Biol.* **2008**, *3*, 180–192.
(66) Ditchfield, C.; Johnson, V. L.; Tighe, A.; Ellston, R.; Haworth, C.; Johnson, T.; Mortlock, A.; Keen, N.; Taylor, S. S. *J. Cell. Biol.* **2003**, *161*, 267–280.
(67) Fancelli, D.; et al. *J. Med. Chem.* **2005**, *48*, 3080–3084.
(68) Harrington, E. A.; Bebbington, D.; Moore, J.; Rasmussen, R. K.; Ajose-Adeogun, A. O.; Nakayama, T.; Graham, J. A.; Demur, C.; Hercend, T.; Diu-Hercend, A.; Su, M.; Golec, J. M. C.; Miller, K. M. *Nat. Med.* **2004**, *10*, 262–267.
(69) Manfredi, M. G.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4106–4111.
(70) Musich, J. A.; Rapoport, H. *J. Am. Chem. Soc.* **1978**, *100*, 4865–4872.